

Graser, Christopher; Fujiwara, Takako; García, Julián; van Veelen, Matthijs

Working Paper

Repeated games with partner choice

Tinbergen Institute Discussion Paper, No. TI 2024-038/I

Provided in Cooperation with:

Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Graser, Christopher; Fujiwara, Takako; García, Julián; van Veelen, Matthijs (2024) : Repeated games with partner choice, Tinbergen Institute Discussion Paper, No. TI 2024-038/I, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/298044>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

TI 2024-038/I
Tinbergen Institute Discussion Paper

Repeated games with partner choice

Christopher Graser¹

Takako Fujiwara-Greve²

Julian García³

Matthijs van Veelen⁴

1 Harvard University

2 Keio University

3 Monash University

4 University of Amsterdam and Tinbergen Institute

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <https://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Repeated games with partner choice

Christopher Graser*, Takako Fujiwara-Greve[†],
Julian García[‡] and Matthijs van Veelen[§]

May 27, 2024

Abstract

Repetition is a classic mechanism for the evolution of cooperation. The standard way to study repeated games is to assume that there is an exogenous probability with which every interaction is repeated. If it is sufficiently likely that interactions are repeated, then reciprocity and cooperation can evolve together in repeated prisoner's dilemmas. Who individuals interact with can however also be under their control, or at least to some degree. If we change the standard model so that it allows for individuals to terminate the interaction with their current partner, and find someone else to play their prisoner's dilemmas with, then this limits the effectiveness of disciplining each other within the partnership, as one can always leave to escape punishment. The option to leave can however also be used to get away from someone who is not cooperating, which also has a disciplining effect. We find that

*Dana-Farber Cancer Institute, Harvard University, Boston, MA 02215, USA

[†]Department of Economics, Keio University, Tokyo, 108-8345 Japan

[‡]Department of Data Science and AI, Monash University, Melbourne, VIC 3800, Australia

[§]Department of Economics, Universiteit van Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, and Tinbergen Institute.

the net effect of introducing the option to leave on cooperation is positive; with the option to leave, the average amount of cooperation that evolves in simulations is substantially higher than without. One of the reasons for this increase in cooperation is that partner choice creates endogenous phenotypic assortment. The model thereby produces a good match with many forms of human cooperation in repeated settings, where we end up interacting, not only with random others that we cannot separate from, once matched, or with others that are genetically related to us, but also with partners that we choose to stay with, and that end up being similarly dependable not to defect on us as we are not to defect on them.

1 Introduction

In the prisoner’s dilemma, repetition can stabilize cooperation. For cooperation to be stable, players need to condition their behavior on the past actions of their interaction partner. If their partner does not cooperate, or does not cooperate enough, then reciprocal players respond with defecting, or with defecting more than they otherwise would. When faced with reciprocal partners, the self-interested thing to do can be to cooperate now in order to receive cooperation in the future. If prisoner’s dilemmas are repeated, this allows for reciprocity and cooperation to evolve together (Axelrod and Hamilton, 1981; Boyd and Lorberbaum, 1987; Fudenberg and Maskin, 1990; Nowak and Sigmund, 1990; Lorberbaum, 1994; Bendor and Swistak, 1995; Binmore and Samuelson, 1992; Cooper, 1996; Volij, 2002; Imhof et al., 2005; van Veelen et al., 2012; Press and Dyson, 2012; Adami and Hintze, 2013; Hilbe et al., 2013; Stewart and Plotkin, 2013, 2014; García and van Veelen, 2016; van Veelen and García, 2019; Dal Bó and Pujals, 2020).

The standard setup in models for the co-evolution of reciprocity and cooperation assumes that randomly matched individuals are tied to their partner

until the repeated game ends. Who plays with whom therefore is determined exogenously. In this paper, we allow for players to end their interaction with their current partner, and look for someone else to continue playing with. Real life interactions are heterogeneous in the degree to which humans are tied to their partners. Some types of interactions allow for easy ways to change partners, others impose higher thresholds for dissolving a partnership, but all interactions find themselves somewhere on the spectrum between the standard setting, where changing partners is not possible at all, and the setting of this paper, where partners can be left, and new partners can be found, after any round of the game.

We also assume that players are not informed about their new partner's past choices. If players are informed about what their partner did in previous interactions with other players, then this could be used to enforce cooperation through norms (Okuno-Fujiwara and Postlewaite, 1995), and it would allow for reputation building (dos Santos et al., 2011, 2013), or indirect reciprocity (Nowak and Sigmund, 1998; Ohtsuki and Iwasa, 2006). By considering a minimal setting in which no information is shared with new partners, we eliminate these possibilities. This way we isolate the role of partner choice in a minimalistic setting, without prior information (cf. Noë and Hammerstein (1994, 1995); McNamara et al. (2008); Barrett and Henzi (2006); Barclay (2013)).

We analyze the evolutionary dynamics in a population playing the repeated prisoner's dilemma, and we investigate whether the option to leave undermines or facilitates the evolution of cooperation. There is a modest number of papers that have a setup in which players have the option to leave (Schuessler, 1989; Vanberg and Congleton, 1992; Enquist and Leimar, 1993; Yamagishi et al., 1994; Kranton, 1996; Ghosh and Ray, 1996; Carmichael and MacLeod, 1997; Sherratt and Roberts, 1998; Hayashi and Yamagishi, 1998; Aktipis, 2004, 2011; Fujiwara-Greve and Okuno-Fujiwara, 2009; Fujiwara-

Greve et al., 2015; Fujiwara-Greve and Okuno-Fujiwara, 2016, 2017; Izquierdo et al., 2010, 2014, 2021; Vesely and Yang, 2010, 2012). While most of these papers do not combine a full game-theoretical analysis with studying the evolutionary dynamics, this literature does contain findings that are relevant. The most important one is a result that states that with the option to leave, there are no equilibria in which all players start cooperating in the first period of every new interaction. The rationale for this is straightforward. Any population in which all individuals do start cooperating right from the beginning can be invaded by a mutant that takes advantage of this, by defecting and leaving after the first period. Such a mutant would get the highest possible payoff in every round, while the resident could at most earn an average payoff equal to the payoff of mutual cooperation. Without the option to leave, there are equilibria with full cooperation, and the fact that fully cooperative equilibria do not exist if leaving is allowed for is a reflection of the downside of the option to leave, which would allow for cheaters to get away with defection and escape punishment whenever cooperation starts in the first round.

To prevent invasions by defect-and-run mutants, a simple solution could be to start every new partnership with a defection. Depending on the parameters of the game, however, one round of mutual defection may not be enough to avoid exploitation. For some combinations of the benefit-to-cost ratio and the continuation probability, starting to cooperate in round 2 may still leave the door open for a mutant that sits out one round of mutual defection, then defects on a resident that starts cooperating in round 2, and subsequently leaves in order to repeat this with its next partner. In Appendix B.3 we specify the minimum length of this initial string of defections, which is also referred to as the *trust-building phase* (Fujiwara-Greve and Okuno-Fujiwara, 2009; Fujiwara-Greve et al., 2015; Fujiwara-Greve and Okuno-Fujiwara, 2016, 2017; Izquierdo et al., 2021). We also translate a result that implies that as

soon as a trust-building phase that meets this threshold is included, cooperative equilibria do exist (Fujiwara-Greve and Okuno-Fujiwara, 2009).

The fact that the option to leave rules out fully cooperative equilibria, and may require multiple periods of trust-building, suggests that leaving might be bad for the evolution of cooperation. Below, we will see that this is not the case, but before we can see how the downside of the option to leave can be more than offset by ways in which it can contribute to the evolution of cooperation, we will have to describe the model we use, both for the theoretical results and for the simulations.

2 The model setup

Individuals are matched to play prisoner’s dilemmas. In this paper, we restrict attention to prisoner’s dilemmas with equal gains from switching (Nowak and Sigmund, 1990), where the cost of cooperating instead of defecting is c , irrespective of whether the opponent cooperates or defects, and the benefits to the other player are b , again irrespective of what the opponent plays herself.

$$\begin{bmatrix} & C & D \\ C & b - c & -c \\ D & b & 0 \end{bmatrix}$$

In order to simplify the notation and analysis further, and without loss of generality, we normalize the costs to $c = 1$. This means that we can interpret the b in the payoff matrix as the benefit-to-cost ratio.

Strategies are represented by finite state automata (FSAs). Figure 2 de-

picts some examples of FSAs. The colours of the states represent the output when the FSA is in this state: red means defect; blue means cooperate; and black means that this FSA terminates the interaction. All FSAs start in the leftmost state when they begin interacting with a new partner, and the arrows indicate to which state the FSA moves in response to their partner's action. After termination, the FSA does not have to transition to any state; it restarts the interaction with its new partner in the leftmost state. Representing strategies as FSAs allows agent-based simulations to explore a very rich and complete space of strategies; any thinkable strategy can be approximated arbitrarily closely by an FSA, and all FSA's can be reached by a finite sequence of mutations.

In the model, we assume that after every round, the pair is broken up exogenously with probability $1 - \delta$, where $\delta \in (0, 1)$. Pairs can also be broken up because one of the players, or both, choose to end the interaction. All broken-up pairs go to the *matching pool*, in which they are re-matched before the subsequent stage game starts. Re-matching happens uniformly at random; all pairs of individuals from the matching pool are equally likely to be formed. The matching pool is not a random draw from the population as a whole. If it would only contain individuals coming from pairs that are broken up exogenously, then the frequencies in the matching pool would match the frequencies in the population as a whole. However, the matching pool also contains individuals that broke up with their partner themselves, and individuals that are broken up with. Whether that happens is determined by the combination of strategies in the pair.

For the theoretical results, we assume an infinitely large population. We also assume that, given the shares of the strategies present in the population, and given the exogenous continuation probability δ , the shares of pairs that consist of the different possible combinations of strategies, and the rounds of the game they are in, are in short-run equilibrium. If there is only one

strategy present in the population, and it never leaves, then it is in principle possible that all pairs are in their first round of play. As pairs are broken up randomly, however, over time the population will converge to a state in which the ratio of pairs that are in the first round of play and pairs that are in their n th round of play is 1 to δ^{n-1} . The intuition for this is that all pairs start in round 1, while the probability for any pair of making it to the n th round is δ^{n-1} . With more than one strategy present, including strategies that may choose to leave, these calculations become more complicated Fujiwara-Greve and Okuno-Fujiwara (2009); Izquierdo et al. (2021). We describe the way of calculating the short-run equilibrium, and the expected payoffs that this short-run equilibrium implies in detail in Appendix A.

3 Long-run equilibrium and indirect invasions

The short-run equilibrium takes the composition of the population as given. In the long run, however, mutation and selection can change the composition of the population. If the population is in a state in which no mutant could ever get an expected payoff that is higher than the expected payoff of the strategies present in the population, and all strategies present in the population have the same expected payoffs, then this population state is a Nash equilibrium. For applying other equilibrium concepts, like evolutionary stability, neutral stability, and robustness against indirect invasions, we also use this separation of time scales, and assume that the population is in short-run equilibrium in order to calculate the expected payoffs of all strategies.

For repeated games without the option to leave, we know that there are many strategies that are neutrally stable (NSS) (Bendor and Swistak, 1995), but no strategies that are robust against indirect invasions (RAII) (van Veelen, 2012; van Veelen et al., 2012; García and van Veelen, 2016). This means that every Nash equilibrium can be invaded indirectly; for every Nash equi-

librium, there is a neutral mutant that, if it goes to fixation, opens the door for a second mutant, that then has a selective advantage. This theoretical result is matched by the fact that in simulations without the option to leave, all Nash equilibria are indeed left in due time, and for reasonably large population sizes, all of those transitions out of equilibria happen through indirect invasions (van Veelen et al., 2012; García and van Veelen, 2016).

With the option to leave, this remains true; all Nash equilibria can be invaded indirectly. The reason is similar to the reason without leaving. For equilibria with positive amounts of cooperation, this cooperation needs to be stabilized with the threat of punishment – which can be to defect (or to defect more than the strategy would otherwise), or to leave. When a population finds itself in such an equilibrium, this punishment is not executed. A mutant that has lost the capacity to punish therefore would be neutral. If random drift allows this neutral mutant to attain a high enough frequency, it would however open the door for a second mutant, that takes advantage of the loss of the capacity to punish (see Appendix B for details).

An equilibrium with defection only can also be invaded indirectly. Without the option to leave, this would require a mutant that *would* cooperate, if its partner initiates it. This is a neutral mutant, and it would open the door for a second mutant that reaps the rewards for initiating cooperation. This stepping stone path requires a minimum δ for the second mutant to have a payoff advantage, and it is still there as a path out of full defection if leaving is allowed for. With leaving, there is also an additional stepping stone path out of fully defecting strategies. A strategy that defects and leaves would be a neutral mutant of any fully defecting strategy. If this strategy takes over the population, it opens the door for a mutant that defects, stays, and cooperates forever after, if it finds its partner has stayed as well. Importantly, this path out does *not* require a minimum δ to constitute an indirect invasion. For low δ , equilibria without any cooperation therefore are less stable with

the option to leave.

With stepping stone paths out of any Nash equilibrium, both with and without the option to leave, the population will visit a variety of equilibria, and it will transition between them through indirect invasions. Which strategies are and which are not Nash equilibria, however, differs between the two settings. For a sufficiently high b and δ , Tit-for-Tat, for example, is an equilibrium without the option to leave, but it is no longer an equilibrium if leaving is allowed for, as it cooperates in the first round. Also for high enough b and δ , a strategy with a trust-building phase, like c_1 in Fig. 2, that, once it reaches the cooperative state, punishes defection with leaving, is an equilibrium when leaving is possible, but this obviously is not a feasible strategy if leaving is not allowed for.

4 Simulations with and without the option to leave

The simulations do not have an infinitely large population. Because it is a simulation, and not a theoretical model, we can moreover not simply assume that it always is in short-run equilibrium. However, even a moderately large finite population tends to be relatively close to short-run equilibrium almost all of the time. More importantly, the long-run dynamics we see in the simulations match what the theory predicts, as we observe sequences of indirect invasions.

The comparison we make here is rather straightforward. In one set of simulations, the output in any state of an FSA can only be cooperate or defect, and not leave; and in the other set of simulations, the output in any state can also be to leave. Without the option to leave, the model then reverts back to the standard model of repeated prisoner's dilemma, with a

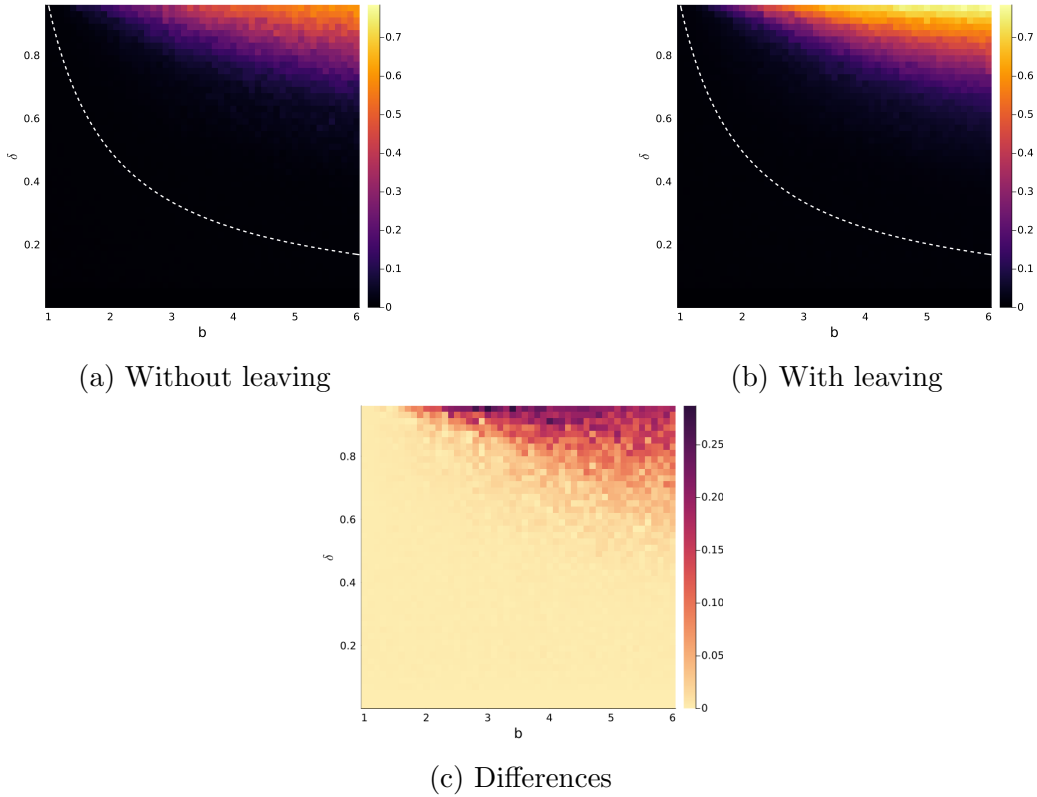


Figure 1: **Simulation results.** Panels (a) and (b) reflect average cooperation levels for a range of benefit-to-cost ratios b and continuation probabilities δ . The population size is $N = 100$. Selection or mutation steps happen at a rate of 0.05 per stage game per matched pair. At a selection or a mutation step, the pair is broken up, and the strategies are replaced by offspring from strategies in the current population, in case of selection, or by a mutant. In expectation, one mutation happens per 250 selection events. The color scale in panels (a) and (b) runs from 0 to 0.785, which is the highest average cooperation level in panel (b). Below the dotted line, no cooperative equilibria exist, both with and without leaving. Panel (c) displays the difference in average cooperation levels between the setting with and the setting without leaving.

continuation probability that is equal to the probability δ with which pairs are not broken up exogenously (see Appendix A.3 for technical details). We

then ran the simulations for a range of b 's, which should be interpreted as the benefit-to-cost ratio, since we normalized the c 's to 1, and a range of δ 's, where $1 - \delta$ is the exogenous breakup probability. Comparing the average amount of cooperation with and without the option to leave, we find that *the option to leave elevates cooperation levels substantially*. For all combinations of b and δ , cooperation is at least as high with leaving as it is without, and the difference is sizable; if we take the average amount of cooperation over the parameter space without leaving – that is: for δ from 0.01 to 0.95 in steps of 0.02, and for b from 1 to 6 in steps of 0.1 – and compare it to the average amount with leaving, then the latter is 42% higher. We should obviously not attach deeper meaning to this exact number, because it is the result of a somewhat arbitrary choice to stop at $b = 6$. If we were to restrict the parameter space to benefit-to-cost ratios between 1 and 5, the difference would be larger than 42%, and if we restrict it to benefit-to-cost ratios between 1 and 7 the gap would be a bit smaller. The number does however justify summarizing the observation that, compared to panel (a) in Fig. 1, in panel (b) the cooperation levels are lifted up across the board, and to a substantial degree.

Allowing for leaving means allowing for more strategies for playing repeated prisoner's dilemmas. FSAs that only have states in which they play cooperate or defect are still included when leaving is made possible; it is just that FSAs in which there are states in which the output is that it leaves are not included in the set of all strategies, when leaving is not allowed for. The increase in size of the set of strategies means that there are more mutants to consider, and for some strategies that are equilibria without the option to leave, that means that they stop being equilibria with leaving, as noted at the end of the previous section. Other strategies remain equilibria, but might nonetheless be left through indirect invasions at a higher or lower rate. In Appendix C, we highlight an example of this by considering the rate at

which fully defecting equilibria are left. On the other hand, extending the set of strategies not only means that there are more mutants to consider, but also more residents that can be equilibria. In the next section we show that in the version of the game with the option to leave, there are indeed new equilibria that punish by leaving. We also show that these equilibria can be more stable than their counterparts that punish by defecting, which also existed in the game without the option to leave.

5 Relative stability of cooperative equilibria with and without leaving

In order to see how the option to leave can add equilibria that are more stable than similar equilibria without leaving, we turn to an example. In this example, we compare an equilibrium strategy that punishes with defection, and one that punishes with leaving. The strategy that punishes with defection is labelled g_1 , and it is best described as Grim Trigger, preceded by a 1-period trust-building phase. The other strategy is c_1 , which also has a 1-period trust-building phase, but responds to defection after the first period by leaving. If b and δ are sufficiently high, both are Nash equilibria – although there is an intermediate part of the parameter space where g_1 is, and c_1 is not (yet) an equilibrium. When playing against copies of themselves, both of these strategies play one round of defection, and then cooperate forever after. The only difference is that one punishes deviations with forever defection, and the other with leaving.

Both strategies can be invaded indirectly in the same way. A strategy that is identical to g_1 , or to c_1 , respectively, but that loses the ability to punish would be a neutral mutant for both. Given that the only difference between g_1 and c_1 is the way they punish, such a neutral mutant ends up being the

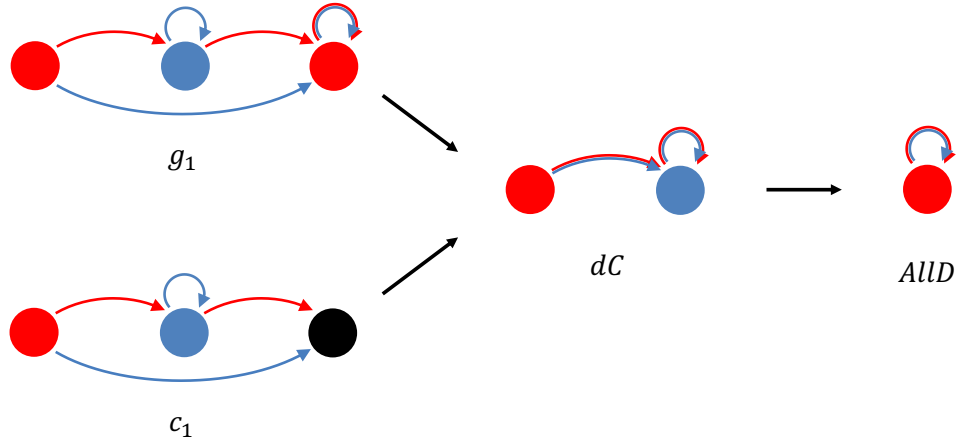


Figure 2: **Indirect invasions in g_1 and c_1 .** The payoffs of g_1 against g_1 , and g_1 against dC are the same as the payoffs of c_1 against c_1 , and c_1 against dC . The payoffs of dC against g_1 and dC against c_1 are also the same. The payoffs of All D against g_1 and All D against c_1 differ, but if dC goes to fixation before All D arises, then both indirect invasions are equally likely to succeed.

exact same strategy for both; this would be dC (see Fig. 2). Strategy dC is neutral, both for g_1 and for c_1 , and therefore it has a fixation probability of $\frac{1}{N}$, where N is the population size. If we can assume that the first mutant has either gone extinct, or gone to fixation, before the next mutant appears, then these indirect invasions into either g_1 or c_1 are equally likely to succeed, because also the second step in the indirect invasion is identical.

If the mutation rate is not low enough to justify this assumption, however, there is a difference. If All D enters the population at a point in time at which both the resident (g_1 or c_1) and the first mutant (dC) are still present, this takes the population to the interior of the simplex (i.e., to a mix of all three strategies). In the interior of the simplex, the replicator dynamics are different, and since the replicator dynamics are also informative about the average dynamics in finite populations, the properties of the finite population

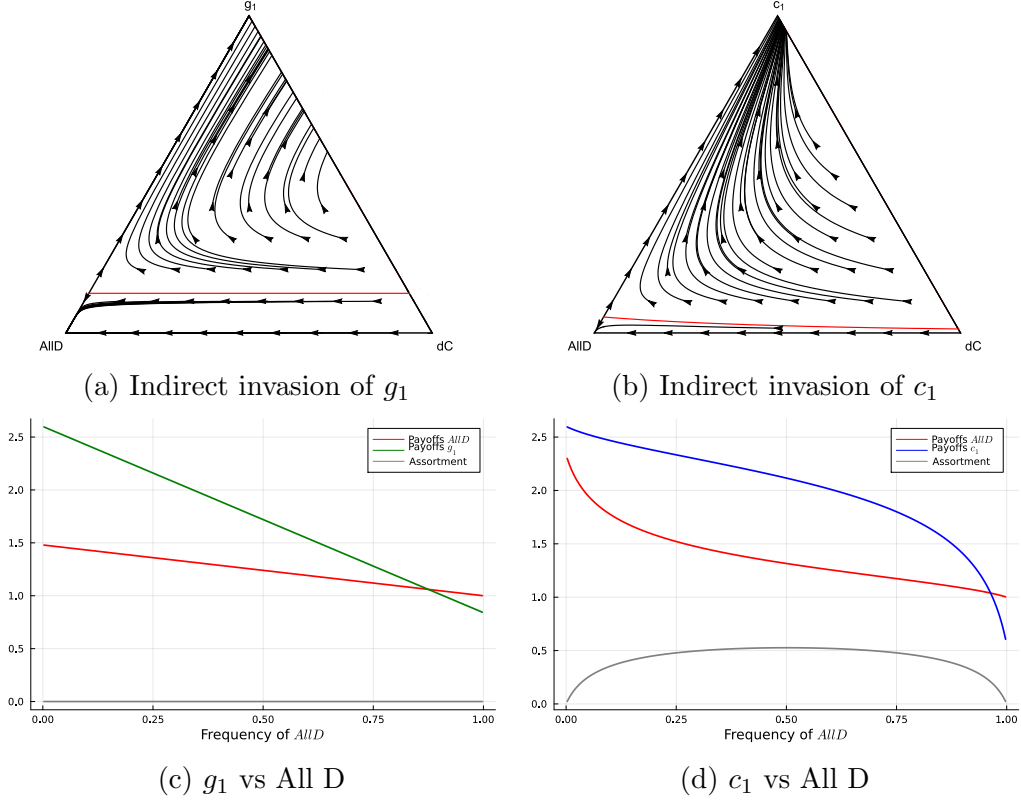


Figure 3: **Replicator dynamics for indirect invasions of g_1 and c_1 .** The red lines in panels (a) and (b) delineate the basins of attraction of All D. Besides the basin of attraction of All D being smaller in panel (b), on trajectories in the interior that do not converge to All D, more dC is weeded out along the way. Panels (c) and (d) provide details relevant to the replicator dynamics on the left edges of the simplex. Gray lines indicate assortment, or relatedness, calculated as the probability with which type i individuals are matched with other type i individuals minus the probability with which an individual of type $j \neq i$ is matched with type i . In other words, $r = P(i|i) - P(i|j)$ for $j \neq i$. Also in the interior of the simplex in panel (b), but not in panel (a), there will be assortment. Payoffs, assortment, and the replicator dynamics are all calculated under the assumption that the distribution of players over pair-types, and over rounds of play, is stationary (Izquierdo et al., 2021). The parameter values used are $b = 3$ and $\delta = 0.8$.

dynamics will be different too. A key ingredient for this difference is that the presence of both c_1 and All D in the same population creates assortment, while with g_1 and All D, this is not the case.

The easier way to see this difference, is to first focus on a population of just g_1 , or just c_1 , respectively, and see what happens if some All D is mixed in. With g_1 and All D, all strategies just stay together, and no leaving happens in any combination. With c_1 and All D, on the other hand, the c_1 -players stick together, while they dissociate from the All D players. At really low frequencies of All D, g_1 actually does a better job at suppressing All D payoffs, because g_1 “binds” the mutant All D’s, and after allowing All D to get a payoff of b once, g_1 then holds All D down to a payoff of 0 in all subsequent periods. Strategy c_1 , on the other hand, cuts All D loose, which allows it to go on and exploit other c_1 ’s. This implies that at the point of invasion (on the very left of panels (c) and (d) of Fig. 3) All D actually gets higher payoffs with c_1 than it does with g_1 . At higher frequencies of All D, however, the assortment that c_1 creates by staying with its own, but dissociating from the All D’s, implies that All D’s mostly find other All D’s in the matching pool. This assortment suppresses the payoffs to those that play All D, when the resident is c_1 . At low frequencies, All D payoffs therefore are suppressed more when the resident is g_1 , while at higher frequencies, All D payoffs are suppressed more when the resident is c_1 . For higher b and δ , the latter effect overpowers the former, making c_1 more stable against All D than g_1 is against All D.

Comparing mixes of g_1 and All D with mixes of c_1 and All D helps understand why there is endogenous assortment with c_1 , and not with g_1 , but the more relevant effect on the stability of the two equilibria we consider here is due to the differences in the interior of the simplex, when three strategies are present. Also in the interior of the simplex, there is no assortment when the strategies present are g_1 , dC , and All D, while there is endogenous assort-

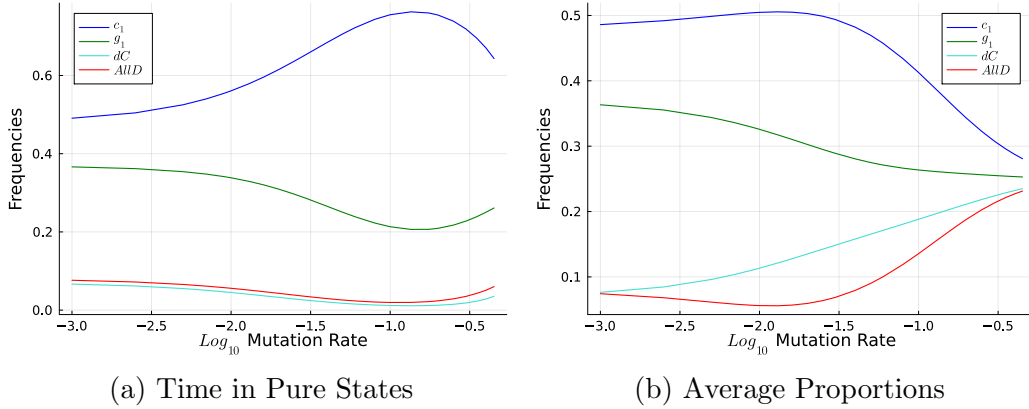


Figure 4: **Relative stability of c_1 and g_1 as a function of the mutation rate.** The plots show properties of the stationary distribution of a Moran process with parameters $N = 40$, $\delta = 0.8$ and $b = 3$ and strategies c_1 , g_1 , dC , and All D. Strategies mutate with equal probabilities into either of the other strategies. The horizontal axis indicates the ratio of mutation steps relative to selection steps. At a mutation rate of -1 , a mutation happens every 10 selection events, and at a mutation rate of -2 , a mutation happens every 100 selection events. Panel (a) shows, out of the time that is spent at monomorphic population states, how much of it is spent at the different strategies respectively. Panel (b) shows the average frequencies of individuals of the respective types.

ment with c_1 , dC , and All D. This implies that trajectories in the interior of the simplex in panel (b) weed out more dC than their counterparts in panel (a) that start at the same points in the interior. The simulations are the noisy, finite population counterpart of these replicator dynamics. That implies that in the finite population dynamics of the simulation program, if the second mutant, All D, appears at the same point on the right edge of the simplex, that is, when the first mutant dC has reached the same frequency after neutrally drifting into the resident (g_1 or c_1), then the dynamic will, in expectation, set this invasion back more if the equilibrium the population started in was c_1 , than it does if the population started in g_1 . That implies that overcoming the tides and currents against this indirect invasion is harder

when c_1 is the resident, and that it takes, on average, more mutations, and therefore more time, to successfully leave c_1 .

We can see the effect of not being in the low-mutation limit by calculating invariant distributions in a finite population for a strategy set that only consists of those four strategies. Increasing the mutation rate increases both the time spent in the c_1 equilibrium relative to the time spent in the g_1 equilibrium, and it increases the average share of c_1 (see Fig. 4). Both of these effects are only reversed when the mutation rate approaches the point where mutation becomes the only ingredient of the dynamics, leaving no room for selection.

6 When punishing with leaving is better than punishing with defecting

The four-strategy model above indicates how equilibria that punish by leaving can be more resistant to indirect invasions compared to equilibria that punish by defecting, away from the low-mutation limit. Using the same set of strategies, but without dC , we can also see how, if leaving is an option, punishing by leaving can outperform punishing by defecting in direct competition between the two modes of punishment. If we focus on a population consisting of strategies c_1 , g_1 , and All D only, then what matters for whether c_1 or g_1 performs better, is the likelihood with which, after breaking up with an All D player, one is re-matched, again, to an All D player. If the probability of trading in one All D partner for another is high, it is better to be g_1 , and sit the current match out. This will result in getting the mutual defection payoff while it lasts, but that then is better than risking wasting another second-round cooperation on a new All D player. If the probability of being matched again with yet another All D player is not too high, on

the other hand, it is better to be c_1 and leave, in the hope of finding a more cooperative partner. The threshold frequency for when the prospect of establishing mutual cooperation makes it worth the risk is favorable for c_1 ; only at very high frequencies of All D is it better to punish with defection (see Fig. 5). Moreover, if punishing by defecting has an advantage over punishing by leaving, both are already losing to All D. Everywhere outside the basin of attraction of All D, where cooperation ends up prevailing, c_1 always beats g_1 (see also Appendix D).

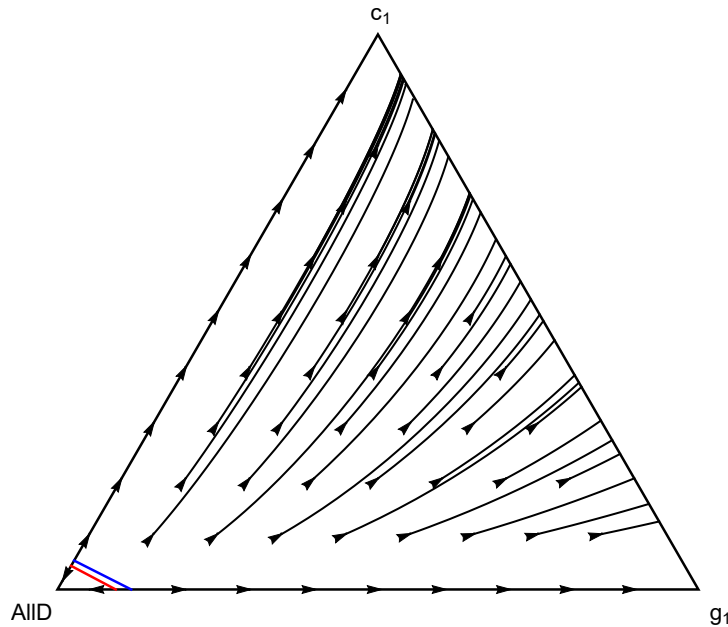


Figure 5: **Leaving or defecting.** Everywhere, except for the area down/left from the red line, c_1 , which punishes with leaving, outperforms g_1 , which punishes with defection. The small area where g_1 outperforms c_1 lies entirely within the basin of attraction of All D, which is delineated by the blue line. Appendix D contains a proof that this holds for all values of b and δ , and also for pairs of cooperative strategies with longer trust-building phases. This implies that, all else equal, if reciprocity evolves, those that punish with leaving always do better than those that punish with defection. The parameter values used for this simplex are $b = 3$ and $\delta = 0.8$.

7 Discussion

The two mechanisms that received the lion share of the attention in the literature on the evolution of cooperation are kin selection – sometimes also classified as population structure – and repetition. Population structure typically refers to any deviation from a well mixed population, in which individuals are matched randomly. This includes interactions on networks (Lieberman et al., 2005; Ohtsuki et al., 2006; Taylor et al., 2007; Santos and Pacheco, 2005; Allen et al., 2017), or within groups (Boyd and Richerson, 1988; Wilson and Wilson, 2007; Traulsen and Nowak, 2006; Simon et al., 2013; Luo, 2014; Akdeniz and van Veelen, 2020). In those models, local dispersal causes neighbouring individuals, or individuals within the same group, to have an increased probability of being identical by descent, and when they do, the mechanism at work is kin selection (Hamilton, 1964a,b; Kay et al., 2020).

Our model relates, first and foremost, to the second mechanism, in which repetition allows for reciprocity and cooperation to co-evolve (Axelrod and Hamilton, 1981; Boyd and Lorberbaum, 1987; Fudenberg and Maskin, 1990; Nowak and Sigmund, 1990; Lorberbaum, 1994; Bendor and Swistak, 1995; Binmore and Samuelson, 1992; Cooper, 1996; Volij, 2002; Imhof et al., 2005; van Veelen et al., 2012; Press and Dyson, 2012; Adami and Hintze, 2013; Hilbe et al., 2013; Stewart and Plotkin, 2013, 2014; García and van Veelen, 2016; van Veelen and García, 2019; Dal Bó and Pujals, 2020). Our version deviates from the standard setup, in that it allows for individuals to leave their current partner, and seek out someone else to play prisoner’s dilemmas with. This shortens the long arm of reciprocity, because it allows individuals to run from punishment by defection. The option to leave however turns out to increase rather than reduce the average amount of cooperation that evolves. By allowing individuals to get up and leave, the model also introduces the possibility of partner choice, and this can create endogenous

assortment in mixed populations. Away from the low mutation limit, this can make equilibria in which defections are punished with leaving more stable than equilibria in which defections are punished with defections. Partner choice therefore seems to be at least as powerful a mechanism for the evolution of cooperation in repeated games. Already in a minimal setting, in which players have no prior information about the partner they are matched with, and only have the interactions within the repeated game itself to base their stay/go decisions on, this mechanism works very well.

Why humans cooperate with their siblings, even if it is costly, or why we are altruistic towards our offspring is well explained by kin selection. The research on the evolution of human cooperation therefore naturally centers around the question why we also cooperate with non-kin. For this we tend to turn towards repeated games and the reciprocity that can evolve there, or to the interaction between repetition and population structure (van Veenen et al., 2012; Efferson et al., 2024). Our model points to the power of a third mechanism, which is partner choice Noë and Hammerstein (1994, 1995); Sherratt and Roberts (1998); Barrett and Henzi (2006); McNamara et al. (2008); Barclay (2013). The assortment that this generates is different from the exogenous assortment that features in kin selection models. The assortment in our model is endogenous, and not based on identity by descent. Individuals in our model stay with their partner purely based on the experienced *phenotype* of their partner, and are not playing with others that are related to them, where relatedness would determine the probability with which they inherited their strategy from the same individual. This phenotypic assortment, where unrelated, similarly dependable cooperators end up playing with each other, may be a better match with the long-lasting cooperation we observe in humans, who tend to exert some influence over who they cooperate with, if they can, and who cooperate with genetically unrelated others.

Code availability

The code used in our simulations is publicly available on Github:

<https://github.com/cjgraser/Repeated-Games-and-Partner-Choice>.

References

- Adami, C., Hintze, A., 2013. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications* 4, 2193.
- Akdeniz, A., van Veelen, M., 2020. The cancellation effect at the group level. *Evolution* 74, 1246–1254.
- Aktipis, C.A., 2004. Know when to walk away: contingent movement and the evolution of cooperation. *Journal of theoretical biology* 231, 249–260.
- Aktipis, C.A., 2011. Is cooperation viable in mobile organisms? Simple Walk Away rule favors the evolution of cooperation in groups. *Evolution and Human Behavior* 32, 263–276.
- Allen, B., Lippner, G., Chen, Y.T., Fotouhi, B., Momeni, N., Yau, S.T., Nowak, M.A., 2017. Evolutionary dynamics on any population structure. *Nature* 544, 227–230.
- Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. *Science* 211, 1390–1396.
- Barclay, P., 2013. Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior* 34, 164–175.
- Barrett, L., Henzi, S.P., 2006. Monkeys, markets and minds: biological markets and primate sociality, in: *Cooperation in primates and humans: Mechanisms and evolution*. Springer, pp. 209–232.

- Bendor, J., Swistak, P., 1995. Types of evolutionary stability and the problem of cooperation. *Proceedings of the National Academy of Sciences* 92, 3596–3600.
- Binmore, K.G., Samuelson, L., 1992. Evolutionary stability in repeated games played by finite automata. *Journal of economic theory* 57, 278–305.
- Boyd, R., Lorberbaum, J.P., 1987. No pure strategy is evolutionarily stable in the repeated prisoner’s dilemma game. *Nature* 327, 58–59.
- Boyd, R., Richerson, P.J., 1988. *Culture and the evolutionary process*. University of Chicago press.
- Carmichael, H.L., MacLeod, W.B., 1997. Gift giving and the evolution of cooperation. *International Economic Review* 38, 485.
- Cooper, D.J., 1996. Supergames played by finite automata with finite costs of complexity in an evolutionary setting. *Journal of Economic Theory* 68, 266–275.
- Dal Bó, P., Pujals, E.R., 2020. The evolutionary robustness of forgiveness and cooperation. Working paper .
- Efferson, C., Bernhard, H., Fischbacher, U., Fehr, E., 2024. Super-additive cooperation. *Nature* , 1–8.
- Enquist, M., Leimar, O., 1993. The evolution of cooperation in mobile organisms. *Animal Behaviour* 45, 747–757.
- Fudenberg, D., Maskin, E., 1990. Evolution and cooperation in noisy repeated games. *American Economic Review* 80, 274–279.

- Fujiwara-Greve, T., Okuno-Fujiwara, M., 2009. Voluntarily separable repeated prisoner's dilemma. *Review of Economic Studies* 76, 993–1021.
- Fujiwara-Greve, T., Okuno-Fujiwara, M., 2016. Diverse behavior patterns in a symmetric society with voluntary partnerships. Working paper .
- Fujiwara-Greve, T., Okuno-Fujiwara, M., 2017. Long-term cooperation and diverse behavior patterns under voluntary partnerships. Working paper .
- Fujiwara-Greve, T., Okuno-Fujiwara, M., Suzuki, N., 2015. Efficiency may improve when defectors exist. *Economic Theory* 60, 423–460.
- García, J., van Veelen, M., 2016. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. *Journal of Economic Theory* 161, 161–189.
- García, J., van Veelen, M., 2018. No strategy can win in the repeated prisoner's dilemma: linking game theory and computer simulations. *Frontiers in Robotics and AI* 5, 102.
- Ghosh, P., Ray, D., 1996. Cooperation in community interaction without information flows. *The Review of Economic Studies* 63, 491–519.
- Hamilton, W.D., 1964a. The genetical evolution of social behaviour. I. *Journal of Theoretical Biology* 7, 1–16.
- Hamilton, W.D., 1964b. The genetical evolution of social behaviour. II. *Journal of Theoretical Biology* 7, 17–52.
- Hayashi, N., Yamagishi, T., 1998. Selective play: Choosing partners in an uncertain world. *Personality and Social Psychology Review* 2, 276–289.
- Hilbe, C., Nowak, M.A., Sigmund, K., 2013. Evolution of extortion in iterated prisoner's dilemma games. *Proceedings of the National Academy of Sciences* 110, 6913–6918.

- Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences* 102, 10797–10800.
- Izquierdo, L.R., Izquierdo, S.S., Vega-Redondo, F., 2010. The option to leave: Conditional dissociation in the evolution of cooperation. *Journal of Theoretical Biology* 267, 76–84.
- Izquierdo, L.R., Izquierdo, S.S., Vega-Redondo, F., 2014. Leave and let leave: A sufficient condition to explain the evolutionary emergence of cooperation. *Journal of Economic Dynamics and Control* 46, 91–113.
- Izquierdo, S.S., Izquierdo, L.R., van Veelen, M., 2021. Repeated games with endogenous separation. Working paper .
- Kay, T., Keller, L., Lehmann, L., 2020. The evolution of altruism and the serial rediscovery of the role of relatedness. *Proceedings of the National Academy of Sciences* 117, 28894–28898.
- Kranton, R.E., 1996. The formation of cooperative relationships. *The Journal of Law, Economics, and Organization* 12, 214–233.
- Lieberman, E., Hauert, C., Nowak, M.A., 2005. Evolutionary dynamics on graphs. *Nature* 433, 312.
- Lorberbaum, J., 1994. No strategy is evolutionarily stable in the repeated prisoner’s dilemma. *Journal of Theoretical Biology* 168, 117–130.
- Luo, S., 2014. A unifying framework reveals key properties of multilevel selection. *Journal of Theoretical Biology* 341, 41–52.
- McNamara, J.M., Barta, Z., Fromhage, L., Houston, A.I., 2008. The coevolution of choosiness and cooperation. *Nature* 451, 189–192.

- Noë, R., Hammerstein, P., 1994. Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behavioral Ecology & Sociobiology* 35, 1–11.
- Noë, R., Hammerstein, P., 1995. Biological markets. *Trends in Ecology & Evolution* 10, 336–339.
- Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the prisoner’s dilemma. *Acta Applicandae Mathematica* 20, 247–265.
- Nowak, M.A., Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. *Nature* 393, 573–577.
- Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A., 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502.
- Ohtsuki, H., Iwasa, Y., 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology* 239, 435–444.
- Okuno-Fujiwara, M., Postlewaite, A., 1995. Social norms and random matching games. *Games and Economic behavior* 9, 79–109.
- Press, W.H., Dyson, F.J., 2012. Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences* 109, 10409–10413.
- Santos, F.C., Pacheco, J.M., 2005. Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters* 95, 098104.

- dos Santos, M., Rankin, D.J., Wedekind, C., 2011. The evolution of punishment through reputation. *Proceedings of the Royal Society B: Biological Sciences* 278, 371–377.
- dos Santos, M., Rankin, D.J., Wedekind, C., 2013. Human cooperation based on punishment reputation. *Evolution* 67, 2446–2450.
- Schuessler, R., 1989. Exit threats and cooperation under anonymity. *Journal of Conflict Resolution* 33, 728–749.
- Selten, R., Hammerstein, P., 1984. Gaps in Harley’s argument on evolutionarily stable learning rules and in the logic of “tit for tat”. *Behavioral and Brain Sciences* 7, 115–116.
- Sherratt, T.N., Roberts, G., 1998. The evolution of generosity and choosiness in cooperative exchanges. *Journal of Theoretical Biology* 193, 167–177.
- Simon, B., Fletcher, J.A., Doebeli, M., 2013. Towards a general theory of group selection. *Evolution* 67, 1561–1572.
- Stewart, A.J., Plotkin, J.B., 2013. From extortion to generosity, evolution in the iterated prisoner’s dilemma. *Proceedings of the National Academy of Sciences* 110, 15348–15353.
- Stewart, A.J., Plotkin, J.B., 2014. Collapse of cooperation in evolving games. *Proceedings of the National Academy of Sciences* 111, 17558–17563.
- Taylor, P.D., Day, T., Wild, G., 2007. Evolution of cooperation in a finite homogeneous graph. *Nature* 447, 469–472.
- Traulsen, A., Nowak, M.A., 2006. Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences* 103, 10952–10955.

- Vanberg, V.J., Congleton, R.D., 1992. Rationality, morality, and exit. *American Political Science Review* 86, 418–431.
- van Veelen, M., 2012. Robustness against indirect invasions. *Games and Economic Behavior* 74, 382–393.
- van Veelen, M., García, J., 2019. In and out of equilibrium II: Evolution in repeated games with discounting and complexity costs. *Games and Economic Behavior* 115, 113–130.
- van Veelen, M., García, J., Rand, D.G., Nowak, M.A., 2012. Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences USA* 109, 9929–9934.
- Vesely, F., Yang, C.L., 2010. On optimal and neutrally stable population equilibrium in voluntary partnership prisoner’s dilemma games. SSRN .
- Vesely, F., Yang, C.L., 2012. Breakup, secret handshake and neutral stability in repeated prisoner’s dilemma with option to leave: A note. SSRN .
- Volij, O., 2002. In defense of DEFECT. *Games and Economic Behavior* 39, 309–321.
- Wilson, D.S., Wilson, E.O., 2007. Rethinking the theoretical foundations of socio-biology. *Quarterly Review of Biology* 82, 327–348.
- Yamagishi, T., Hayashi, N., Jin, N., 1994. Prisoner’s dilemma networks: selection strategy versus action strategy, in: *Social dilemmas and cooperation*, Springer. pp. 233–250.

A The model

We mostly follow the model setup in (Fujiwara-Greve and Okuno-Fujiwara, 2009; Fujiwara-Greve et al., 2015; Fujiwara-Greve and Okuno-Fujiwara, 2016, 2017). In these papers, the authors consider a repeated prisoner’s dilemma with the option to leave (or, as they call it, with voluntary separation). In (Fujiwara-Greve and Okuno-Fujiwara, 2009) the stage game payoffs are given by four parameters. We chose to restrict attention to prisoner’s dilemmas with equal gains from switching, which means that we have only two parameters, if we also normalize the payoff of mutual defection to 0.

$$\begin{bmatrix} & C & D \\ C & b - c & -c \\ D & b & 0 \end{bmatrix}$$

Without loss of generality, we then chose $c = 1$, which is equivalent to dividing all entries by c . This implies that we interpret b as the benefit-to-cost ratio. This is the payoff matrix we use in the Main Text. Adding 1 to all entries does not change the equilibrium analysis, or the replicator dynamics, and because the simulations require non-negative payoffs, in order to be consistent with the payoff matrix used in the simulations, here in the Appendix we use the following equivalent payoff matrix.

$$\begin{bmatrix} & C & D \\ C & b & 0 \\ D & b + 1 & 1 \end{bmatrix}$$

Assuming equal gains from switching limits attention to prisoner’s dilemmas in which the cost of cooperation to oneself, and the benefits to the other, are independent of whether the other cooperates or defects. This is a real restriction. Normalizing the lowest payoff to 0, and setting the cost of cooperation to 1, on the other hand, is not a restriction of generality.

The model furthermore assumes that all pairs have an exogenous breakup probability of $1 - \delta$, with $\delta \in (0, 1)$. Pairs can also be broken up because one of the players, or both, choose to end the partnership. All individuals from broken-up pairs are re-matched before the subsequent stage game, so there are no costs associated with having to wait to play again. Re-matching happens uniformly at random among the set of all unmatched individuals between two stage games. We refer to this set as the *matching pool*.

The model in Fujiwara-Greve and Okuno-Fujiwara (2009) is slightly different, in that the δ there represents the probability of an individual surviving between two stage games. If an individual dies, then the pair it was in is broken up. The dead individual is replaced by a new one, and both the newborn individual and the surviving individual from the pair go to the matching pool. This translates to an exogenous breakup probability of $1 - \delta^2$. If the newborn individual simply inherits the strategy of the deceased individual, there is no difference between, on the one hand, surviving and, on the other hand, dying and being replaced. In that case, the model from (Fujiwara-Greve and Okuno-Fujiwara, 2009; Fujiwara-Greve et al., 2015; Fujiwara-Greve and Okuno-Fujiwara, 2016, 2017) and the model in this paper are equivalent, up to the change from δ^2 to δ representing the probability for a pair not to be broken up exogenously. The latter allows for a bit shorter notation and representation of the results. Death and birth can also be reinterpreted as update events, at which individuals may switch to different strategies, based on the payoffs these strategies get in the population.

Introducing selection in (Fujiwara-Greve and Okuno-Fujiwara, 2009; Fujiwara-Greve et al., 2015; Fujiwara-Greve and Okuno-Fujiwara, 2016, 2017) would imply that the offspring does not simply inherit the strategy of the deceased individual, or that only the current shares of the different strategies present in the population determine the probabilities with which they reproduce, but that also the payoffs of individuals in the current population determine the probabilities with which they pass on their strategy. In our model, introducing selection requires introducing replacement or update events altogether. We will introduce selection explicitly below, in the section that describes the simulations. All replacement events there will lead to the pair splitting up, but not all splitting up of pairs will be the result of replacement events. Before we do this, however, we describe how the payoffs are calculated for a given composition of the population.

A.1 Calculating payoffs

In repeated games without the option to leave, the expected payoff of a strategy can be computed by first computing the discounted payoffs for the strategy against every strategy present in the population, and then weighing those payoffs according to the shares of those different strategies in the population. In repeated games with the option to leave, this is no longer possible. Given that players can choose to leave, also in the absence of exogenous breakups, not all matches last equally long. That implies that who is matched to whom becomes partly endogenous, as some combinations of strategies stick together longer than others.

In order to be able to calculate the payoffs, we will have to assume that a strategy distribution is in a steady state, that is, it is in short-run equilibrium. Below, we will first give the formula for the average payoff of a strategy, and then we will show why the assumption that the population is in a steady

state implies that the formula should have this form. In general, the repeated prisoner's dilemma has infinitely many pure strategies, with or without the option to leave. However, the populations we consider will always be finite mixtures of pure strategies. We will therefore refer to the strategies as s_i , for $i = 1, 2, \dots, n$. This allows us to denote the frequencies of these strategies *in the matching pool* by a vector, x , where x_i is the frequency of strategy s_i in the matching pool.

For a combination of strategies s_i and s_j , T_{ij} will denote the number of periods that s_i and s_j play, if no exogenous breakup occurs. The expected duration of this match, given that the pair is broken up exogenously with probability $1 - \delta$ after every round, is

$$L_{ij} = 1 + \delta + \dots + \delta^{T_{ij}-1} = \sum_{t=0}^{T_{ij}-1} \delta^t = \frac{1 - \delta^{T_{ij}}}{1 - \delta}.$$

In order to compute the expected payoff V_{ij} of s_i against s_j , let a_t^{ij} denote the action profile at period t as strategies s_i and s_j meet, for $t = 1, \dots, T_{ij}$. The payoff to strategy s_i , matched with strategy s_j , in round t is denoted by $u(a_t^{ij})$. Then,

$$V_{ij} = \sum_{t=0}^{T_{ij}-1} \delta^t u(a_t^{ij}).$$

With x_i representing the stationary frequency of strategy s_i in the matching pool, the average per-period payoff of strategy s_i is given by

$$v_i(x) = \frac{\sum_{j=1}^n x_j V_{ij}}{\sum_{j=1}^n x_j L_{ij}}. \quad (1)$$

A.2 Why these are the average payoffs, if we assume short-run equilibrium

In short-run equilibrium, the shares of given pairs of strategies that find themselves in any given round should be constant. We denote the fraction of ordered pairs *in the population as a whole*, that consist of an s_i -player and an s_j -player, and that find themselves in the t 'th round of their interaction, by $q_{ij,t}$. Because the game is symmetric, $q_{ij,t} = q_{ji,t}$ for any pair of strategies (s_i, s_j) and any t . It is also possible to derive the average payoff of an s_i -player using unordered pairs, but it will be more straightforward to think of ordered pairs.

In order for the frequencies $q_{ij,t}$ to be constant in the short-run dynamics, the share of (s_i, s_j) -pairs that find themselves in the t 'th round of their interaction should be δ times the share of (s_i, s_j) -pairs that find themselves in the $t - 1$ 'st round of their interaction – provided that $t \leq T_{ij}$. The reason is that between one period and the next, the probability that a group is not broken up exogenously is δ . All groups of (s_i, s_j) -pairs that find themselves in the $t - 1$ 'st round of the game, and that are not broken up exogenously, become groups of (s_i, s_j) -pairs that find themselves in the t 'th round – unless $t > T_{ij}$. For $q_{ij,t}$ to be constant in the short-run dynamics (see Izquierdo et al. (2021)), it therefore has to be equal to $\delta q_{ij,t-1}$, while $q_{ij,t} = 0$ if $t > T_{ij}$.

Because no strategy can end a partnership before the first round, the shares in the first round moreover must be proportional to the shares in the matching pool, if we assume short-run equilibrium. For ordered pairs, that means that $q_{ij,1}$ should be proportional to x_i times x_j .

If we then choose

$$q_{ij,t} = \frac{x_i x_j \delta^{t-1}}{\sum_{k=1}^n \sum_{l=1}^n \sum_{u=1}^{T_{kl}} x_k x_l \delta^{u-1}} \text{ for } 1 \leq t \leq T_{ij}$$

and

$$q_{ij,t} = 0 \text{ for } t > T_{ij},$$

then they are proper fractions, as they add up to 1, and all of these restrictions for them to be in short-run equilibrium are satisfied. These therefore are the shares of different types of (ordered) pairs that find themselves at different times in their interaction in short-run equilibrium.

Given that the population is in short-run equilibrium, the average payoff to a strategy in the population as a whole, in in any given round, is constant over time. The average payoff to strategy s_i in the population for which x represents the frequencies of the strategies in the matching pool is given by

$$\begin{aligned} v_i(x) &= \frac{\sum_{j=1}^n \sum_{t=1}^{T_{ij}} q_{ij,t} u(a_t^{i,j})}{\sum_{j=1}^n \sum_{t=1}^{T_{ij}} q_{i,j,t}} = \\ &= \frac{\sum_{j=1}^n \sum_{t=1}^{T_{ij}} x_i x_j \delta^{t-1} u(a_t^{i,j})}{\sum_{j=1}^n \sum_{t=1}^{T_{ij}} x_i x_j \delta^{t-1}} = \\ &= \frac{x_i \sum_{j=1}^n x_j \sum_{t=0}^{T_{ij}-1} \delta^t u(a_t^{i,j})}{x_i \sum_{j=1}^n x_j \sum_{t=0}^{T_{ij}-1} \delta^t} = \\ &= \frac{\sum_{j=1}^n x_j V_{ij}}{\sum_{j=1}^n x_j L_{ij}}. \end{aligned}$$

This is formula [1] above.

A.3 With or without leaving

If we then were to revert back to a setting in which leaving is not allowed, then this would imply that we only allow for pairs of strategies s_i and s_j with $T_{ij} = \infty$, $L_{ij} = \frac{1}{1-\delta}$, and $V_{ij} = \sum_{t=0}^{\infty} \delta^t u(a_t^{i,j})$. Without leaving, and using the fact that x_i 's are frequencies in the matching pool that add up to one, we find that the average per-period payoff becomes

$$\begin{aligned}
v_i(x) &= \frac{\sum_{j=1}^n x_j V_{ij}}{\sum_{j=1}^n x_j L_{ij}} \\
&= (1 - \delta) \frac{\sum_{j=1}^n x_j V_{ij}}{\sum_{j=1}^n x_j} \\
&= (1 - \delta) \sum_{j=1}^n x_j V_{ij}.
\end{aligned}$$

This is equal to the normalized, discounted payoffs of strategy s_i against the mix of strategies s_j , for $j = 1, \dots, n$ where the discount factor is taken to be the probability δ for a pair not to be broken up exogenously. Calculating a stream of payoffs over time, and then normalizing therefore coincides with taking the average in the population – where normalizing means multiplying with $1 - \delta$. Therefore, in the setup that we created to accommodate for the possibility of leaving, if we choose not to allow for leaving, we are back in a situation that is equivalent to the standard setting without leaving. This also justifies comparing the cases with and without leaving within one and the same framework.

A.4 Frequencies in the matching pool and in the population as a whole

Let us go back to the repeated prisoner’s dilemma with the option to leave. If we want to know how large a share of the population as a whole is playing strategy s_i , we can aggregate the frequencies of pair types over all strategies s_j that s_i could be matched with, and over all periods that a pair can be in. The shares of strategies s_i , in the population as a whole are given by a vector, x^* , where x_i^* is the frequency of strategy s_i in the population as a

whole, for $i = 1, \dots, n$.

$$\begin{aligned}
x_i^* &= \sum_{j=1}^n \sum_{t=1}^{T_{ij}} q_{ij,t} \\
&= \sum_{j=1}^n \sum_{t=1}^{T_{ij}} \frac{x_i x_j \delta^{t-1}}{\sum_{k=1}^n \sum_{l=1}^n \sum_{u=1}^{T_{kl}} x_k x_l \delta^{u-1}} \\
&= \frac{x_i \sum_{j=1}^n x_j \sum_{t=1}^{T_{ij}} \delta^{t-1}}{\sum_{k=1}^n \sum_{l=1}^n x_k x_l \sum_{u=1}^{T_{kl}} \delta^{u-1}} \\
&= \frac{x_i \sum_{j=1}^n x_j L_{ij}}{\sum_{k=1}^n \sum_{l=1}^n x_k x_l L_{kl}}
\end{aligned}$$

Frequencies in the matching pool therefore translate to frequencies in the population as a whole in a relatively straightforward way. Going in the other direction is not straightforward at all; for computing the frequencies in the matching pool from the frequencies in the population as a whole, there is generally no closed form expression; see Izquierdo et al. (2021). The fact that the frequencies in the population as a whole and the frequencies in the matching pool can be different is part of what makes the game with the option to leave interesting, as it is an unavoidable symptom of the endogenous assortment that the setup with leaving allows for. This does however complicate the appropriate definition of the replicator dynamics, and it also makes the definitions of a Nash equilibrium and a Neutrally Stable State a bit more involved; see Izquierdo et al. (2021). Theorem 1 in Izquierdo et al. (2021) moreover shows that a Neutrally Stable State, according to the definition in the same paper, is Lyapunov stable in the replicator dynamics. This is not the case for a Neutrally Stable Distribution as defined in Fujiwara-Greve and Okuno-Fujiwara (2009).

A.5 Simulations

Translating this framework into a finite population model that we can simulate to analyze long-run dynamics requires two further ingredients; we need to specify a selection procedure, and a mutation procedure.

To a large degree, we follow the approach taken in van Veelen et al. (2012); García and van Veelen (2016, 2018), and extend it to the setting with endogenous separation. As in those papers, strategies are represented by finite state automata (FSAs) in our simulations. We do however depart from the model in those papers, in that we allow for the output in every state of the automaton – except for the starting state – to also take on the value *leave*, besides *cooperate* and *defect*. The starting state cannot have the output *leave*, as every strategy has to specify a stage game action for the first stage game after it is being matched. A formal definition of an FSA is given below, after the section that formally defines histories and strategies. For understanding how they work, it is enough to be able to read their graphical representations in for instance Fig. 6.

Because the option to leave is not present in van Veelen et al. (2012); García and van Veelen (2016, 2018), they chose for all pairs in the population to start playing the repeated prisoner’s dilemma at the same moment in time. In that setup, when all pairs have finished their repeated interactions (with independent, random durations), the entire population is updated based on the payoffs the individuals earned in the repeated game. This makes the selection dynamics a Wright-Fisher process. With leaving and re-matching, the beginnings of the repeated interactions are not synchronised, and therefore it will be hard to use the Wright-Fisher process. Here we use a Moran process instead, which implies that in each selection step, only one individual is replaced by the offspring of a random member of the population, where the probability for any individual to be the parent is proportional to their payoff. Selection steps in our simulations occur between stage games, and

the relevant payoffs for reproduction are the payoffs from the last stage game. The pair to which the individual who gets replaced belonged is broken up, and both the offspring that replaces the partner that is chosen to be replaced, and the partner that is not replaced join the matching pool. Similarly, if one member of a matched pair mutates, which in our simulations also happens between stage games, and independently of selection events, the corresponding pair is broken up, and both the mutant and its former partner join the matching pool. In this paper, the presence or absence of the option to leave is the only difference between the settings we compare. The simulation model without leaving in this paper is therefore different from the simulation model in van Veelen et al. (2012); García and van Veelen (2016, 2018), because here we do not have synchronised beginnings of the repeated games, and we have a Moran process instead of a Wright-Fisher process. The simulations in van Veelen and García (2019) also use a Moran process, but because they assume the low mutation limit, and because leaving is not an option there, they can calculate the fixation probabilities explicitly, rather than running agent-based simulations.

The exogenous break-ups that our selection and mutation processes bring about are, of course, relevant for the probability with which matched pairs are broken up in our simulations. To keep consistency with the basic theoretical framework, when talking about the simulations, δ will denote the probability with which the pair is not broken up exogenously. The remaining $1 - \delta$ is the probability that at least one individual in the pair mutates, plus the the probability that at least one member in the pair dies and is replaced in a selection event, plus an additional probability that the pair is broken up without mutation or selection events. This last probability is equal to the difference between $1 - \delta$ and the sum of the other two probabilities. In the simulations, the probability of mutation and selection events therefore put an upper bound on the continuation probability, and in order to be able

to go to relatively high continuation probabilities, we need to accept that replacement and mutation events are relatively rare.

The mutation procedure we chose is a modified version of the mutation procedure in van Veelen et al. (2012); García and van Veelen (2016, 2018). In their procedure, all mutations make single small changes; they change the output in a state, delete or add a state, or change a transition. We chose a mutation procedure that also allows for multiple changes in one go. The reason to do this, is that with leaving, mutants that would have a selective advantage are sometimes multiple changes away from the resident automaton. In order not to have populations remain in disequilibrium states, because the mutation procedure cannot find the mutants that would have an advantage, we therefore allow for multiple changes in one mutation event.

In the process of developing this mutation procedure we explored various different specifications. Since a mutation that redirects a link, or removes a state, can render one or more states unreachable, there are different options how to deal with those unreachable states. We tried out mutation procedures that after a mutation 1) remove all unreachable states; 2) reconnect the unreachable states; and 3) keep the unreachable states around, which means that future mutations may reconnect to this “junk DNA”. We also tried different relative probabilities for adding or deleting states (see the code for details). All of these modifications affect the average size and complexity of the automata that evolve, and the share of unreached and neutrally-evolving states in the automata. We did however find the same qualitative results presented in this paper for all these versions of the mutation procedure. The simulations reported in this paper all use the third mutation procedure.

A.6 Histories and strategies

Here, we give a formal definition of the game. This is not needed to understand the main text, or the main ideas here in the Appendix, but for the proofs of the theorems, a bit of additional formality can be useful.

We assume a set of players $I = \{1, 2\}$, an action space $A = \{C, D\}$, equal for both players, and a payoff function $u : A \times A \rightarrow \mathbb{R}$. A history at time t is a list of actions played up to and including time $t - 1$, where an empty pair of brackets is used to denote the null history. If $a_{t,i}$ is the action played by player i at time t , then these histories are:

$$\begin{aligned} h_1 &= () \\ h_t &= ((a_{1,1}, a_{1,2}), \dots, (a_{t-1,1}, a_{t-1,2})), \quad t = 2, 3, \dots \end{aligned}$$

Sometimes we will also write $(h_t, (a_{t,1}, a_{t,2}))$ for a history h_{t+1} . We will also write h_t^\leftarrow for history h_t , as seen from the perspective of player 2. The set of possible histories at time t is:

$$\begin{aligned} H_1 &= \{h_1\} \\ H_t &= \prod_{i=1}^{t-1} (A \times A) \quad t = 2, 3, \dots \end{aligned}$$

and the set of all possible histories is:

$$H = \bigcup_{t=1}^{\infty} H_t.$$

A pure strategy is a function $s : H \rightarrow A$ that maps histories to the action space.

A.7 Finite state automata

For the simulations, we will restrict attention to strategies that can be represented by finite state automata (FSAs). An FSA, or a Moore machine M , is a tuple $\{\{1, \dots, n_M\}, \lambda_M, \mu_M\}$, where n_M is the number of states, $\lambda_M : \{1, \dots, n_M\} \rightarrow \{C, D\}$ gives the output in every state, and $\mu_M : \{1, \dots, n_M\} \times \{C, D\} \rightarrow \{1, \dots, n_M\}$ gives the transitions as a function of the state, and the action of the other player. Sometimes the formal definition of a machine also specifies in which state the machine starts, but because the states can always be renumbered so that the starting state is the first, we assume, without loss of generality, that the machine starts in state 1. Sometimes we will also order the remaining states, so that, if it plays against a copy of itself, then, in the absence of errors, it will transition from state i to state $i + 1$, until for the first time it goes back to a state it has already been in, or remains in the state it is currently in. Figure 6 gives a graphical representation of an FSA.

B Theoretical results

B.1 No ESS

There is no finite mix of strategies that is evolutionarily stable in the repeated prisoner's dilemma, with or without the option to leave. This is caused by the fact that in any population, there are histories that are never reached. This implies that it is always possible to construct a neutral mutant, that behaves identically to the resident after all histories that are reached, and differently in at least one history that is not; see also Selten and Hammerstein (1984) and Proposition 6 in García and van Veelen (2016).

B.2 No strategy that is RAI

For any cooperative equilibrium, strategies that behave identically to the resident, but do not punish, can drift into the population as neutral mutants. Once such a neutral mutant has established itself as the new resident, it can subsequently be invaded by less cooperative strategies. Such indirect invasions are possible, whether punishment happens with defection or with leaving. For repeated games without the option to leave, this is observed in Theorem 7 of García and van Veelen (2016). Here we give the equivalent result for the case with leaving.

Theorem 1. *No pure strategy that, when playing a copy of itself, ever plays C in the repeated prisoner's dilemma with endogenous separation is RAI.*

Proof. If a strategy cooperates against itself after some history h_t , it must disincentivise defecting at this history, either by leaving or by defecting in some future round. Otherwise, a mutant that defects at this history earns higher payoffs, and could invade directly. However, as the history in which a player defects after h_t is never reached, any strategy that does not punish

this defection and otherwise behaves identically to the resident is a neutral mutant. Such a strategy could then be invaded directly by a strategy which behaves identically to the resident but defects at h_t . \square

If the option to leave is not there, and a strategy does not always cooperate when it meets a copy of itself, then indirect invasions with an increase in cooperation are also feasible, provided that δ is sufficiently high. This is observed in Theorem 8 of García and van Veelen (2016). The indirect invasion there is a neutral mutant that would cooperate (more), if the other initiates it, but does not initiate (additional) cooperation itself, followed by a mutation that initiates (additional) cooperation. For the second mutant to have an advantage, the δ needs to be high enough.

With leaving, here we present a simpler result, which is that full defection can always be invaded. It is simpler, in the sense that the starting point is a population in which all strategies defect with all other strategies, while the result for the case without leaving takes as a starting point all strategies that fall short of full cooperation. On the other hand, while without leaving, there is a restriction on the δ for this indirect invasion to exist, with leaving there is no such restriction, and an indirect invasion exists regardless of the value of δ .

Theorem 2. *No mixture of strategies, in which all strategies always defect with all strategies present in the population, is RAI.*

Proof. In a population in which every strategy present never cooperates with any other strategy present, all strategies are also defecting in the first period. The strategy that defects and leaves in the first period (which we will denote with d_0) therefore is a neutral mutant of any population with universal mutual defection. If by random drift d_0 takes over, and becomes the new resident, then a strategy that defects in the first round, and then stays, and cooperates forever after, earns a payoff of 1 against d_0 , as does d_0 against

d_0 , and as does d_0 against this strategy, while this strategy earns a payoff of $1 + \delta(b - 1) > 1$ against itself. Any mix with universal defection therefore can be invaded indirectly. \square

Theorem 1 and 2 combined imply that, also if leaving is allowed for, there is no pure strategy that is RAI. These are the equivalents of Theorems 7 and 8 in García and van Veelen (2016), which do the same for the case without leaving. The same paper also contains extensions of these two theorems to finite mixtures of strategies. A similar extension of Theorem 1 above can be made for the case with leaving. Theorem 2 already includes mixtures as it is, and does not need extending.

Given that no strategy exists that is RAI, the best we can do, is find strategies that are NSS. This is what we will look for below.

B.3 Pure strategies with a trust building phase

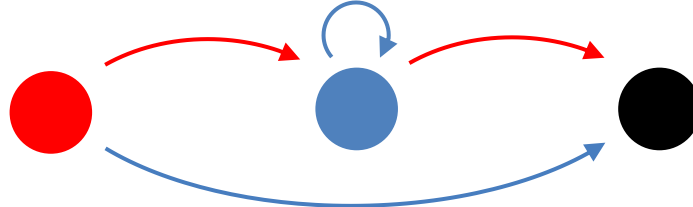


Figure 6: **Strategy c_1** . As a finite state automaton, c_1 has three states; in state 1, the output is D (red); in state 2 the output is C (blue); and in state 3, the output is to leave (black). The arrows indicate to which state the strategy goes, after observing an action (C or D) by its opponent. It does not have to transition to any state after termination by either player.

In this section we translate a result from Fujiwara-Greve and Okuno-Fujiwara (2009) to fit our simplified model with equal gains from switching,

and an exogenous breakup rate rather than an individual death rate. Strategies c_T , that cooperate after a trust building phase of length T , are defined in Fujiwara-Greve and Okuno-Fujiwara (2009) as follows.

Definition 3. For any $T \in \mathbb{N}_0$ let c_T be a strategy that

- plays D in period t if $t \leq T$,
- plays stay after round $t = 1, \dots, T$ if and only if (D, D) is observed in round t ,
- plays C in period t if $t \geq T + 1$,
- plays stay after round $t = T + 1, T + 2, \dots$ if and only if (C, C) is observed in round t .

Note that rounds of play in Fujiwara-Greve and Okuno-Fujiwara (2009), and in this paper, are indexed starting at $t = 1$, so that strategy c_0 has no trust building phase. Also, it is possible to give alternative, but equivalent definitions of a trust building strategy c_T – where equivalent means that the alternative definition prescribes the same behaviour, not just when matched with c_T , but when matched with any possible strategy. Alternative definitions would specify actions differently only for histories that simply cannot occur, given that one plays according to c_T oneself.

In the simulations, strategies are encoded as FSAs that only respond to actions by the other player. This restricts their response to a history of length t ending in (D, D) to be the same as the response to the same history, but now ending in (C, D) . Note, however, that, since output is deterministic, only one of them can occur, which makes responses to histories that cannot occur entirely inconsequential. Although the FSA representation in Figure 6, for example, is not the exact same strategy as c_1 , as defined in Fujiwara-Greve and Okuno-Fujiwara (2009), it will play exactly the same as c_1 , defined above, when matched with any possible strategy.

Because our stage game is a bit simpler than the stage game in Fujiwara-Greve and Okuno-Fujiwara (2009), and because we have an exogenous breakup probability, rather than a probability for an individual to die, we can also give a simpler version of their Proposition 1, with a simpler proof. Our simpler version does however include an explicit formula for the thresholds over which c_T 's become equilibria. These are depicted in Figure 7. Our version also includes the straightforward extension that, *below* the threshold, no equilibria exist that start cooperating before round $T + 1$.

Theorem 4. (*Fujiwara-Greve & Okuno-Fujiwara, 2009*)

1. c_T is a Nash equilibrium if $b \geq \frac{1-\delta^{T+1}}{\delta-\delta^{T+1}}$.
2. All equilibria start with at least $T + 1$ rounds of defection if $b < \frac{1-\delta^{T+1}}{\delta-\delta^{T+1}}$.

Proof. Any other strategy s' that does not do the same against c_T as c_T does against c_T would have to deviate by either

1. playing leave if (D, D) is observed in some round $K \leq T$,
2. playing C in some round $K \leq T$
3. playing D in some round $K \geq T + 1$, or
4. playing leave if (C, C) is observed in some round $K \geq T + 1$.

Round K is the *first* round in which s' does something that is different from what c_T would do. All of these deviations can then be combined with additional, inconsequential changes beyond period K .

In order to be able to compare the payoffs of these alternative strategies, we first calculate the average per-period payoff of c_T in a population where

everybody is playing c_T .

$$\begin{aligned} u(c_T, c_T) &= \frac{1 + \delta \cdot 1 + \dots + \delta^{T-1} \cdot 1 + \delta^T \cdot b + \delta^{T+1} \cdot b + \dots}{1 + \delta + \delta^2 + \dots} = (1 - \delta) \frac{(1 - \delta^T) \cdot 1}{1 - \delta} + (1 - \delta) \frac{\delta^T \cdot b}{1 - \delta} \\ &= 1 - \delta^T + \delta^T b = 1 + \delta^T (b - 1) \end{aligned}$$

1. In a population in which everybody (else) plays c_T , the average per period payoff of a strategy s^1 that leaves in round $K \leq T$ is

$$u(s^1, c_T) = 1 < 1 + \delta^T (b - 1) = u(c_T, c_T).$$

2. In a population in which everybody (else) plays c_T , the average per period payoff of a strategy s^2 , that plays cooperate in round $K \leq T$, is

$$u(s^2, c_T) = 1 - (1 - \delta) \frac{\delta^{K-1}}{1 - \delta^K} < 1 < u(c_T, c_T).$$

This deviation therefore is worse than the deviation by s^1 , which makes perfect sense, because playing C and being left is always worse than playing D and leaving.

3. In a population in which everybody (else) plays c_T , the average per period payoff of a strategy s^3 that plays D in round $K \geq T + 1$ is

$$\begin{aligned} u(s^3, c_T) &= \frac{1 - \delta^T}{1 - \delta^K} + \frac{\delta^T - \delta^{K-1}}{1 - \delta^K} b + \frac{\delta^{K-1} - \delta^K}{1 - \delta^K} (b + 1) \\ &= \frac{1 - \delta^T}{1 - \delta^K} + \frac{\delta^T - \delta^K}{1 - \delta^K} b + \frac{\delta^{K-1} - \delta^K}{1 - \delta^K}. \end{aligned}$$

For the comparison with $u(c_T, c_T)$, it may be even easier to rewrite this in a way that reflects that we can also consider the first T periods separately,

and then consider the repeating pattern starting from period $T + 1$:

$$\begin{aligned} u(s^3, c_T) &= 1 - \delta^T + \frac{\delta^T - \delta^K}{1 - \delta^K} b + \frac{\delta^{K-1} - \delta^K}{1 - \delta^K} + \frac{\delta^K - \delta^{K+T}}{1 - \delta^K} \\ &= 1 - \delta^T + \frac{\delta^T - \delta^K}{1 - \delta^K} b + \frac{\delta^{K-1} - \delta^{K+T}}{1 - \delta^K}. \end{aligned}$$

Therefore, $u(c_T, c_T) \geq u(s^3, c_T)$ if and only if

$$\begin{aligned} 1 - \delta^T + \delta^T b &\geq 1 - \delta^T + \frac{\delta^T - \delta^K}{1 - \delta^K} b + \frac{\delta^{K-1} - \delta^{K+T}}{1 - \delta^K} \\ \delta^T b &\geq \frac{\delta^T - \delta^K}{1 - \delta^K} b + \frac{\delta^{K-1} - \delta^{K+T}}{1 - \delta^K} \\ \delta^T (1 - \delta^K) b &\geq (\delta^T - \delta^K) b + \delta^{K-1} - \delta^{K+T} \\ (\delta^K - \delta^{K+T}) b &\geq \delta^{K-1} - \delta^{K+T} \\ (\delta - \delta^{T+1}) b &\geq 1 - \delta^{T+1} \\ b &\geq \frac{1 - \delta^{T+1}}{\delta - \delta^{T+1}}. \end{aligned}$$

4. Because $u(s^3, c_T) > u(s^4, c_T)$ for obvious reasons (playing D and being left is always better than playing C and leaving), and because $u(s^2, c_T) < u(s^1, c_T) < u(c_T, c_T)$ for all combinations of δ and T , as seen above, this condition now guarantees that no other strategy can do better against c_T than c_T . This proves the first part of the theorem.

For the second part of the theorem, we define d_T as the strategy that plays defect for the first $T + 1$ rounds, and leaves after round $T + 1$. Let s be as a strategy that, when playing against a copy of itself, plays defect in the first T rounds, and plays cooperate for the first time in round $T + 1$. The same calculations that gave us the threshold above imply that, if $b < \frac{1 - \delta^{T+1}}{\delta - \delta^{T+1}}$, then $u(d_T, c_T) > u(c_T, c_T)$, and because $u(d_T, s) = u(d_T, c_T)$ and $u(c_T, c_T) \geq u(s, s)$, this implies that also $u(d_T, s) > u(s, s)$. Moreover, if $b < \frac{1 - \delta^k}{\delta - \delta^k}$ for

$k \geq 3$, then also $b < \frac{1-\delta^{k-1}}{\delta-\delta^{k-1}}$, so strategies with even shorter trust-building phases can also be invaded. \square

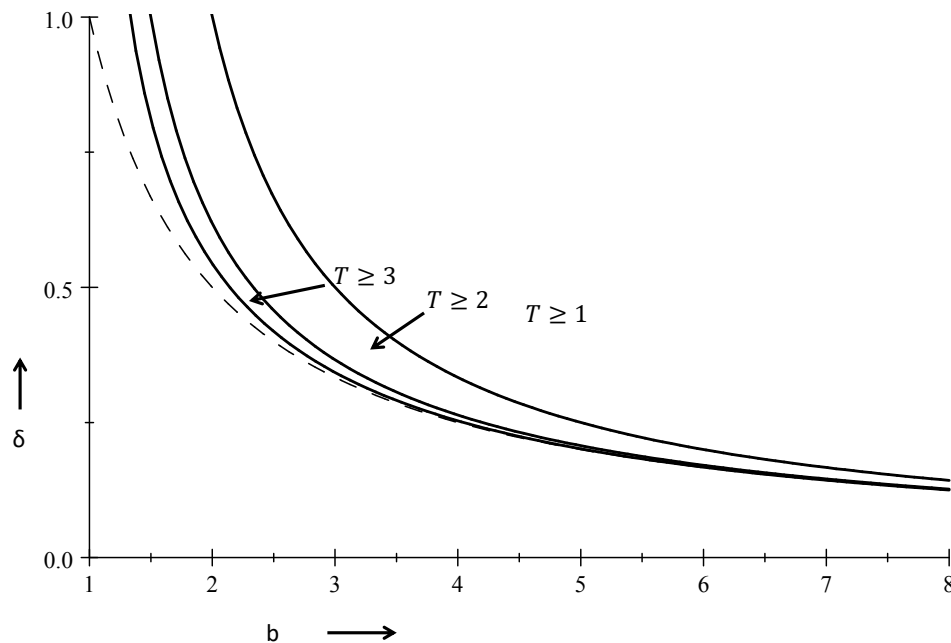


Figure 7: **Pure c_T equilibria: NSS, not RAI.** Right/up from the black lines is where different trust building strategies c_T constitute equilibria. For any point right/up from the dotted line, given by $b = \frac{1}{\delta}$, there is some T such that c_T is a Nash equilibrium, as are all trust building strategies with longer trust building phases. Proposition 2 in Fujiwara-Greve and Okuno-Fujiwara (2009), appropriately adapted, implies that these equilibria are also NSS.

Also Proposition 2 in Fujiwara-Greve and Okuno-Fujiwara (2009) allows for a simpler version, given our simpler stage game. Besides translating it to the simpler setting, we also switch from the definition of neutral stability

in Fujiwara-Greve and Okuno-Fujiwara (2009) to the definition of neutral stability in Izquierdo et al. (2021). For pure equilibria, the latter definition is more restrictive; the definition in Fujiwara-Greve and Okuno-Fujiwara (2009) requires that no pure mutant can invade, while the definition in Izquierdo et al. (2021) also requires mixtures of strategies not being able to invade. Theorem 1 in Izquierdo et al. (2021) shows that being neutrally stable in the latter definition moreover implies Lyapunov stability in the replicator dynamics. Our Theorem 5 therefore is less general than Proposition 2 in Fujiwara-Greve and Okuno-Fujiwara (2009), but it is also more specific about the bounds, and it shows that c_T satisfies a slightly more restrictive, and dynamically more relevant equilibrium concept.

Theorem 5. (*Fujiwara-Greve & Okuno-Fujiwara, 2009*)

$$c_T \text{ is an NSS if } b > \frac{1-\delta^{T+1}}{\delta-\delta^{T+1}}$$

Proof. Here, we use Lemma 1 from Izquierdo et al. (2021). This lemma states that if a strategy i never leaves when playing against itself, and $u(s_i, s_i) > u(s_j, s_i)$ for any strategy s_j such that a different history unfolds when strategy s_j plays with strategy s_i , then strategy s_i is a neutrally stable strategy. Using this lemma, we can then re-use the proof of Theorem 2, with strict inequalities for strategy s^3 . \square

Although c_T is an NSS when the condition in Theorem 5 is satisfied, it is, by Theorem 1, not RAI; there is always an indirect invasion with loss of cooperation.

Indirect invasions with an increase in cooperation are sometimes also feasible. A first mutant would play like c_T does, except that it would choose to stay whenever the other player cooperates, also in the first T rounds, and moreover that it would respond to the other player cooperating in any of those earlier rounds by also switching to playing C . The second mutant then

could be c_K with $K < T - 1$, which implies it initiates cooperation at least 2 rounds before T . This would earn a payoff of $1 - \delta^K + \delta^{K+1}b$ against the first mutant, while $v(c_T, c_T) = 1 - \delta^T + \delta^T b$, which is also the payoff of the first mutant against itself. The second mutant would therefore have an advantage over the first if

$$b > \frac{1 - \delta^{T-K}}{\delta - \delta^{T-K}}.$$

The mutant with the largest advantage would be c_0 , which starts cooperating immediately in round 1. Even though c_0 would itself not be a Nash equilibrium, it remains true that if any c_K can invade indirectly, then so can c_0 . The condition for there to be a $K < T - 1$ such that c_K can invade c_T indirectly, therefore, is that c_0 can invade indirectly. This condition is $b > \frac{1 - \delta^T}{\delta - \delta^T}$, which is the same as the condition for c_{T-1} to be an equilibrium. In Figure 7, this implies that between the thresholds for c_T and c_{T-1} , c_T is an equilibrium, and while it can be invaded indirectly by a sequence that undermines cooperation, it cannot be invaded indirectly by a sequence that increases cooperation, without undermining it first. Between the threshold for c_{T-1} and c_T , one could therefore argue that c_T is more stable than the other trust building strategies.

C Getting away from All D at a higher rate

The dynamics, both with and without the option to leave, tend to go through similar phases. A population state in which there is no cooperation whatsoever is invaded indirectly, after which the population settles on a cooperative equilibrium. A subsequent indirect invasion then takes it back to a fully defecting equilibrium, such as, for example, All D. Sometimes an indirect invasion will take the population from one equilibrium with a positive amount

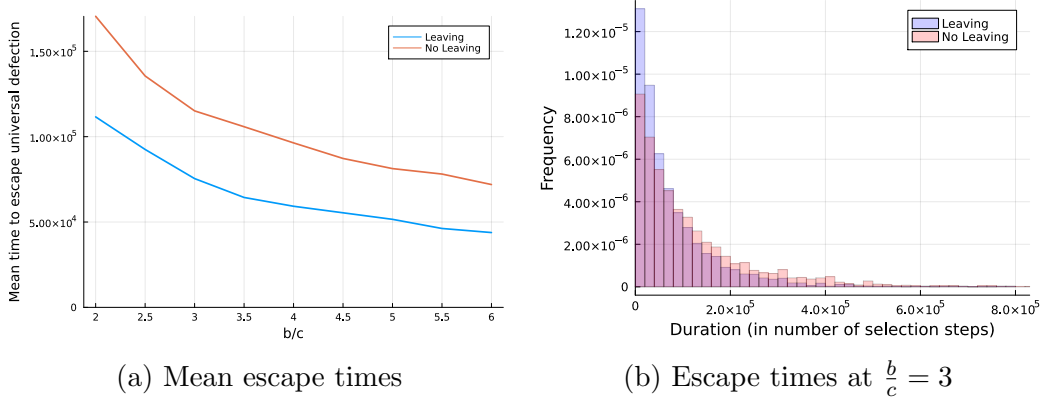


Figure 8: **Stability of universal defection with and without the option to leave.** Panel (a) shows, for a variety of b/c -ratios, how much time (measured as the number of consecutive selection steps) the simulated population spends on average with a resident that only defects, before a more cooperative mutant invades successfully. The continuation probability is fixed at $\delta = 0.9$. Panel (b) shows the distribution of escape times for a b/c -ratio of 3. The arrival rate of mutants is the same, with or without the option to leave; it is just the sets of mutants, and the distributions over those, that are different between the two settings.

of cooperation to another one, with a different amount of cooperation, but transitions with a complete loss of cooperation are sufficiently more frequent to ensure that the population returns to the set of fully defecting equilibria very regularly. Given that cooperation tends to break down completely, before it is re-established, both with and without the option to leave, any change in the rate at which states like All D are left is consequential.

As is illustrated in Figure 8, the average time it takes for mutation and selection to find a path out of equilibria that are equivalent to All D is lower in the setting with the option to leave than it is without it. This also contributes to the fact that there is more cooperation in the game with the option to leave than there is without, even though the option to leave limits the effectiveness of punishment with defection.

D Punishing with leaving is always selected for when cooperation is selected for

Assume the following three strategies. Strategy 1 is g_n , which is Grim Trigger, preceded by a trust-building phase of n periods. In an FSA, this trust-building phase is represented by n states in which the output is defection, and in which the FSA always moves to the next state. After having gone through the n trust-building states, a cooperative state is reached, in which g_n stays as long as the opponent plays C , and a D state that is absorbing, and where g_n transitions to if the opponent played D (see Figure 2 in the main text for g_1). Strategy 2 is c_n , which is essentially the same strategy, but instead of punishing defections after the first n rounds with defecting forever, it leaves (see also the definition in Section S2 above). Strategy 3 is All D. If x_1 , x_2 , and x_3 represent the shares of these three strategies in the matching pool, then their payoffs are:

$$\begin{aligned}
 v_1(x) &= \frac{\left(\frac{1-\delta^n}{1-\delta} + \frac{\delta^n b}{1-\delta}\right) x_1 + \left(\frac{1-\delta^n}{1-\delta} + \frac{\delta^n b}{1-\delta}\right) x_2 + \left(\frac{1-\delta^n}{1-\delta} + \frac{\delta^{n+1}}{1-\delta}\right) x_3}{\frac{1}{1-\delta} x_1 + \frac{1}{1-\delta} x_2 + \frac{1}{1-\delta} x_3} \\
 &= 1 + \delta^n (b - 1) - \delta^n (b - \delta) x_3 \\
 v_2(x) &= \frac{\left(\frac{1-\delta^n}{1-\delta} + \frac{\delta^n b}{1-\delta}\right) x_1 + \left(\frac{1-\delta^n}{1-\delta} + \frac{\delta^n b}{1-\delta}\right) x_2 + \left(\frac{1-\delta^n}{1-\delta}\right) x_3}{\frac{1}{1-\delta} x_1 + \frac{1}{1-\delta} x_2 + \frac{1-\delta^{n+1}}{1-\delta} x_3} \\
 &= \frac{1 - \delta^n + \delta^n b (1 - x_3)}{1 - \delta^{n+1} x_3} \\
 v_3(x) &= \frac{\left(\frac{1}{1-\delta} + \delta^n b\right) x_1 + \left(\frac{1-\delta^{n+1}}{1-\delta} + \delta^n b\right) x_2 + \left(\frac{1}{1-\delta}\right) x_3}{\frac{1}{1-\delta} x_1 + \frac{1-\delta^{n+1}}{1-\delta} x_2 + \frac{1}{1-\delta} x_3} \\
 &= \frac{1 + \delta^n (1 - \delta) b (x_1 + x_2) - \delta^{n+1} x_2}{1 - \delta^{n+1} x_2}.
 \end{aligned}$$

Strategies 1 and 2 do equally well if $x_3 = \frac{b-1}{b-\delta}$; if we fill in $x_3 = \frac{b-1}{b-\delta}$ in the formulas for the payoffs above, we find that $v_1(x) = v_2(x) = 1$. From the formula for $v_3(x)$ above, we can immediately see that $v_3 \geq 1$, and that $v_3 > 1$ if $x_1 + x_2 > 0$, which holds because $x_3 = \frac{b-1}{b-\delta} < 1$, since $\delta < 1$. This implies that no trajectory of the replicator dynamics that starts at a population state x with $x_3 \geq \frac{b-1}{b-\delta}$ can ever reach a population state with $x_3 < \frac{b-1}{b-\delta}$. For that to happen, the trajectory would have to pass through some population state for which $x_3 = \frac{b-1}{b-\delta}$ at which the share of strategy 3 is decreasing. This would contradict the fact that the payoff of strategy 3 is higher than the payoff of the other strategies for all population states x for which $x_3 = \frac{b-1}{b-\delta}$.

The x here reflects shares in the matching pool. Section S1 above details how these translate to shares in the population as a whole. Points on the simplices in the main text all depict shares of strategies in the population as a whole. The line where $x_3 = \frac{b-1}{b-\delta}$ here does have a constant share of strategy 3 in the matching pool, but that does not imply that the share of strategy 3 in the population as a whole is also constant. If the simplex in Fig 5 would have represented shares in the matching pool, the red line would have been parallel to the edge of the simplex at which $x_3 = 0$, but since the simplex represents shares in the population as a whole, it is not.